

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-143451

(43)Date of publication of application : 25.05.2001

(51)Int.Cl. G11B 27/10
G06F 17/30
G10L 11/02
G10L 15/04
G10L 15/10
G10L 15/00
G11B 27/36
H04N 5/91

(21)Application number : 11-326445

(71)Applicant : NIPPON HOSO KYOKAI <NHK>

(22)Date of filing : 17.11.1999

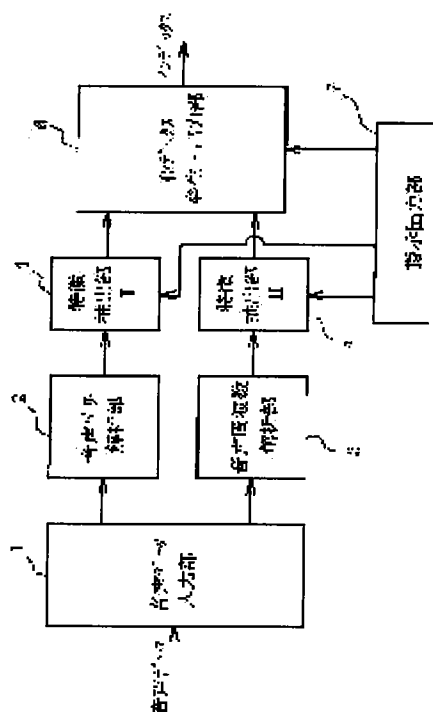
(72)Inventor : SANO MASAKI
INOUE SEIKI
SUMIYOSHI HIDEKI

(54) AUTOMATIC INDEX GENERATING DEVICE AND AUTOMATIC INDEX APPLYING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To solve the problem in which an automatic index generating device suited to a purpose does not exist conventionally, even though efficient retrieval is enabled in the case of producing or secondarily utilizing a digest program by automatically generating and applying an index to be shown in the excited part of video of a sports event or the like.

SOLUTION: This device is provided with respective analysis parts 2 and 3 for analyzing the power level and frequency of audio data, respective feature extracting parts 4 and 5 for extracting features from the respective power level and frequency and a part 6 for generating/outputting the index from the provided features.



LEGAL STATUS

[Date of request for examination] 28.01.2005

[Date of sending the examiner's decision of rejection] 30.05.2006

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-143451

(P2001-143451A)

(43) 公開日 平成13年5月25日 (2001.5.25)

(51) Int.Cl. ⁷	識別記号	F I	テームト* (参考)
G 1 1 B 27/10		G 1 1 B 27/10	A 5 B 0 7 5
G 0 6 F 17/30		27/36	A 5 C 0 5 3
G 1 0 L 11/02		G 0 6 F 15/40	3 7 0 D 5 D 0 1 5
15/04		15/401	3 1 0 A 5 D 0 7 7
15/10		G 1 0 L 3/00	5 1 3 A 5 D 0 8 7
審査請求 未請求 請求項の数 2 O L (全 5 頁) 最終頁に続く			

(21) 出願番号 特願平11-326445

(22) 出願日 平成11年11月17日 (1999. 11. 17)

(71) 出願人 000004352

日本放送協会

東京都渋谷区神南2丁目2番1号

(72) 発明者 佐野 雅規

東京都世田谷区砧1丁目10番11号 日本放送協会 放送技術研究所内

(72) 発明者 井上 誠喜

東京都世田谷区砧1丁目10番11号 日本放送協会 放送技術研究所内

(74) 代理人 100059258

弁理士 杉村 暁秀 (外2名)

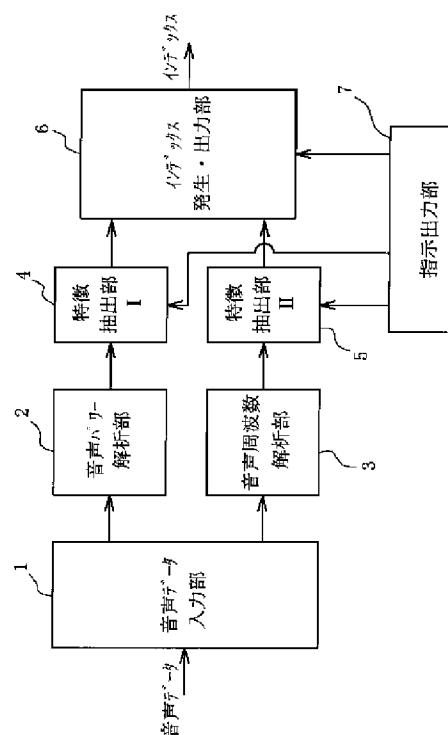
最終頁に続く

(54) 【発明の名称】 自動インデックス発生装置ならびにインデックス付与装置

(57) 【要約】

【課題】 スポーツイベントなどの映像において、盛り上がった部分などでそれを示すインデックスを自動的に発生させ、かつ付与することができれば、ダイジェスト番組の制作や2次利用の際の効率的な検索が可能になるが、従来、このような目的に合った自動インデックス発生装置は存在しなかった。

【解決手段】 音声データのパワーレベルと周波数を解析する各解析部2、3と、それぞれからの特徴を抽出する各特徴抽出部4、5と、得られた特徴からインデックスを発生・出力する部分6とを具えて構成した。



【特許請求の範囲】

【請求項1】 音声信号を用いて、イベント中に発生する各種特徴事項に対応するインデックスを自動的に発生する装置であって、
入力音声信号の全周波数帯に亘るパワーレベルを解析する音声パワー解析部と、
該音声パワー解析部による解析結果に少なくとも1つの予め設定したパワーレベルの音声が含まれているか否かを示す第1の特徴を抽出する第1の特徴抽出部と、
前記入力音声信号の周波数および該周波数毎のパワーレベルを解析する音声周波数解析部と、
該音声周波数解析部による解析結果に少なくとも1つの予め設定した周波数および該周波数毎のパワーレベルの音声が含まれているか否かを示す第2の特徴を抽出する第2の特徴抽出部と、
前記第1の特徴抽出部で抽出された第1の特徴と、前記第2の特徴抽出部で抽出された第2の特徴とを組み合わせる前記インデックスを発生するインデックス発生部とを具えてなることを特徴とする自動インデックス発生装置。

【請求項2】 映像・音声記録装置と請求項1記載の自動インデックス発生装置とが具えられ、前記映像・音声記録装置で記録される映像・音声に、前記自動インデックス発生装置により発生したインデックスを付与するように構成したことを特徴とするインデックス付与装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声信号を用いて、イベント中に発生する各種特徴事項に対応するインデックスを自動的に発生する自動インデックス発生装置、およびその装置により発生したインデックスを記録映像・音声に付与するインデックス付与装置に関する。

【0002】

【従来の技術】スポーツイベントなどの映像を撮影する場合、盛り上がった部分で自動的にインデックスを発生させ、さらには、その発生したインデックスをVTRなどで記録すべき映像・音声に多重するなどの形態で付与することができれば、これを利用して全撮影映像からダイジェスト版を制作したり、2次利用の際に特定の場面の検索が可能になるなど極めて便利である。

【0003】この分野で利用できる従来のインデックスとしては、従来、映像信号を用いて発生させた物理的インデックス（被写体の色、形、動きなどに基づいて発生させたもの）と意味的インデックス（登場人物、動作イベントなどに基づいて発生させたもの）とが存在していた。

【0004】また、音声信号を用いてインデックスを発生させるものとして、従来、インタビュー部分、話者特定などを用いてインデックスを発生した例はあるものの、インデックス発生のための装置が複雑になったり、

また、発生したインデックス自体の信頼性にも問題があった。

【0005】

【発明が解決しようとする課題】上述したように、スポーツイベントなどの映像で、盛り上がった部分などにそれを示すインデックスを自動的に発生させ、かつ付与することができれば、ダイジェスト番組の制作や2次利用の際の効率的な検索が可能になる。自動的にインデックスを付与することができない場合、人が映像（全撮影映像）を最初から見ることで、盛り上がりの部分を探し当てたり、また、上記発生させた物理・意味的インデックスを利用して盛り上がったであろう部分を推測するしかない。

【0006】本発明の目的は、上述したような種々の課題を含まない、音声信号を用いてインデックスを発生させる新規な自動インデックス発生装置、ならびにその発生装置で発生したインデックスを記録映像・音声に付与するインデックス付与装置を提供することにある。

【0007】

【課題を解決するための手段】本発明は、イベント会場の音声を入力とし、そのパワーレベルと周波数解析を行って得られた特徴から、例えば、盛り上がった部分や審判が笛を吹いた部分などを検出し、これをインデックスとして利用するようにしたものである。すなわち、本発明自動インデックス発生装置は、音声信号を用いて、イベント中に発生する各種特徴事項に対応するインデックスを自動的に発生する装置であって、入力音声信号の全周波数帯に亘るパワーレベルを解析する音声パワー解析部と、該音声パワー解析部による解析結果に少なくとも1つの予め設定したパワーレベルの音声が含まれているか否かを示す第1の特徴を抽出する第1の特徴抽出部と、前記入力音声信号の周波数および該周波数毎のパワーレベルを解析する音声周波数解析部と、該音声周波数解析部による解析結果に少なくとも1つの予め設定した周波数および該周波数毎のパワーレベルの音声が含まれているか否かを示す第2の特徴を抽出する第2の特徴抽出部と、前記第1の特徴抽出部で抽出された第1の特徴と、前記第2の特徴抽出部で抽出された第2の特徴とを組み合わせる前記インデックスを発生するインデックス発生部とを具えてなることを特徴とするものである。

【0008】また、本発明インデックス付与装置は、映像・音声記録装置と本発明による自動インデックス発生装置とが具えられ、前記映像・音声記録装置で記録される映像・音声に、前記自動インデックス発生装置により発生したインデックスを付与するように構成したことを特徴とするものである。

【0009】

【発明の実施の形態】以下に添付図面を参照し、発明の実施の形態に基づいて本発明を詳細に説明する。上述したように、本発明自動インデックス発生装置ならびにイ

ンデックス付与装置は、音声信号を用いて、スポーツイベントなどの盛り上がった部分で自動的にインデックスを発生させ、また付与するものであるが、これによれば、イベントの流れに対応した各種インデックスを自動的に発生させ、その発生したインデックスを記録映像・音声に付与することを可能にし、それによりイベント内での盛り上がった部分を自動的に抽出することが可能となり、また、イベントにおける特徴をもつ音成分を抽出してそれぞれコード化し、もしくはこれらを組み合わせることで新たな特徴コードを生成し、ダイジェストの作成や映像検索を容易に行えるようにして、情報検索機能を一層向上させることができるようにしたものである。

【0010】以下では、スポーツイベントとしてサッカーを中継放送する場合を例にとりて、インデックスを自動的に発生する本発明自動インデックス発生装置について説明する。図1は、本発明自動インデックス発生装置の一実施形態をブロック図にて示している。図1において、まず、サッカースタジアムのスタンド上部で収音された会場全体の音声である音声データ（信号処理がデジタルの形態で行われるため、音声信号と言わないで音声データと言う、以下同じ）を音声データ入力部1に入力する。この音声データ入力部1にはバッファメモリが用意され、そのメモリに、次段の音声パワー解析部2および音声周波数解析部3におけるパワーレベルと周波数の解析に必要なまとまった音声データを保持するとともに、その保持した音声データをそれら両解析部2および3へ出力することができるようになっている。

【0011】音声パワー解析部2においては、音声データ入力部1から入力された一定時間長の音声データについて、瞬時振幅値を2乗して平方根をとることで、パワーレベルを計算して出力する。また、音声周波数解析部3においては、その周波数解析部に入力される音声データに窓関数を用いて一定時間（例えば、0.05秒）を切り出し、切り出した部分についての周波数解析を行い、例えば、10Hz毎のパワーレベルを計算する。このような計算を、一定時間毎（例えば、0.025秒毎）に行い、その結果を出力する。

【0012】符号4で示す特徴抽出部Iでは、音声パワー解析部2から入力される全周波数帯に亘るパワーレベルのうち、ある一定以上のパワーレベルがある部分（時間帯）を抽出（特徴抽出）して出力する。また、符号5で示す特徴抽出部IIでは、音声周波数解析部3から入力される周波数毎のパワーレベルについて、例えば、審判の笛の音（約3,500Hz付近の周波数成分）や拡声器による応援、会場全体に聞こえるような口笛などの特殊な応援（約1,300Hz付近の周波数成分）の聞こえる時間帯を抽出（特徴抽出）して出力する。

【0013】このように、特徴抽出部Iおよび特徴抽出部II（それぞれ、符号4および5で示される）において、特徴抽出部Iでは全周波数帯に亘るパワーレベル

（例えば、「大」「小」）が、特徴抽出部IIでは周波数毎のパワーレベル（例えば、「大」「小」）がそれぞれ抽出されることになる。例えば、特徴抽出部Iの出力がレベル「大」である場合であって、特徴抽出部IIの3,500Hzの出力のみがレベル「大」であるときは審判の笛の音Sであると認識し、3,500Hz以外の周波数成分の出力もレベル「大」であるときは盛り上がりの音声Mであると認識する。また、特徴抽出部Iの出力がレベル「小」である場合であって、特徴抽出部IIの1,300Hzの出力のみがレベル「大」であるときは口笛などの特殊な応援の音声Cであると認識する。表1は、各種の入力音声についてこの認識をする際に必要となる全周波数帯に亘るパワーレベルの「大」「小」と、周波数毎のパワーレベルの「大」「小」を表にして示すものである。

【0014】

【表1】

全周波数帯に亘るパワーレベル 周波数毎のパワーレベル	全周波数帯に亘るパワーレベルの「大」「小」と、周波数毎のパワーレベルの「大」「小」		
		「大」	「小」
f a [3,500 Hz]	「大」	S, M	
	「小」		C
f b [1,300 Hz]	「大」	M	C
	「小」	S	

全周波数帯に亘るパワーレベルの「大」「小」と、周波数毎のパワーレベルの「大」「小」

(S: 審判の笛、 M: 盛り上がり、 C: 特殊な応援)

【0015】特徴抽出部Iおよび特徴抽出部IIによってそれぞれ抽出された音声の特徴を示す信号は、インデックス発生・出力部6に入力される。インデックス発生・出力部6は、当該部に入力された上記のような各種信号を組み合わせてインデックス（インデックス信号）を発生するよう構成されているが、これは、それらが上記のような多種類の音声であっても、それぞれの音声の特徴を抽出することによりそれらを区別して別のインデックスとして出力することができる。実際には、特徴抽出部Iの出力と特徴抽出部IIの出力との間で各種論理演算を行うことで実行される。例えば、全周波数帯に亘るパワーレベル「大」と3,500Hzのパワーレベル「大」と1,300Hzのパワーレベル「小」との論理積演算を行うことによって審判の笛の音Sを得るなどはその一例である（表1参照）。

【0016】また、指示出力部7は、イベントの種類、インデックスを発生させたい対象となる音声等に対応させて、符号4, 5で示す特徴抽出部I, IIに音声パワーレベルの「大」「小」切り分けの境界の指示、符号5で示す特徴抽出部IIに抽出する周波数の指示、およびインデックス発生・出力部6に両特徴抽出部から抽出されたそれぞれの特徴の組み合わせ方（論理演算の仕方を含む）

等の指示をするためのものである。

【0017】インデックス発生・出力部6から最終的に取り出されるインデックスの一例を図2に示す。ここでも、符号Sは審判の笛の音を、Mは盛り上がりの音声を、また、Cは特殊な応援の音声をそれぞれ示し、かつ、それらが時間軸上で何時発生したかを示している。なお、審判の笛の音の例からも分かるように、笛の音はある時間帯にわたって聞こえるが、インデックスはそれがあることを示す単一のパルス状のフラグであると考えればよい。

【0018】本発明は、以上説明した自動インデックス装置により発生したインデックスを、VTRなどの映像・音声記録再生装置に記録のために供給される映像または音声信号に多重するなどして、何らかの形で付与するインデックス付与装置を含むものである。図3は、このインデックス付与装置によりインデックスが付与された映像のみを取り出してサッカー番組のダイジェスト版を制作する場合を示している。

【0019】図3において、(a)はインデックス、(b)は映像、(c)は制作されたダイジェスト番組である。ここでは、サッカー競技において、全体の番組から符号Mで示す盛り上がりの場面の映像（例えば、盛り上がった部分の前10秒、後5秒を含む映像）のみを取り出して接続することにより、ダイジェスト番組を制作する例を示している。勿論、盛り上がりの場面の映像を全部つながなくても、それら映像をダイジェスト映像の候補とし、その中からさらに絞り込むことにより、一層コンパクトにダイジェスト化された番組を制作することができる。これにより、はじめから全体の映像を見なくても、必要とする部分を絞りこむことができ、番組制作の効率を大幅に向上させることができる。

【0020】以上は、スポーツイベントとしてサッカーをとりあげて説明したが、本発明の対象はこれに限定されるものではなく、例えば、アメリカンフットボール、

バスケットボールなどほかのスポーツ、また、多くの観衆が存在するスポーツ以外のイベント、例えば、コンサートやサーカスなどにも適用可能であること勿論である。

【0021】

【発明の効果】本発明によれば、観衆が多く集まるイベントなどの映像に、音声データを利用してイベントの流れに各種インデックスを自動発生することで、ダイジェスト番組の制作支援、2次利用の際の検索が可能となり、また、スポーツ映像においては、試合の流れを記述する情報ともなり、より高度で統合された番組制作に関する情報管理が可能となることが期待される。

【0022】また、本発明によれば、音声データを解析することで盛り上がった部分を検出でき、同時に特定の音声データ（審判の笛の音など）を検出することでインデックスを発生させ、イベントの流れを記述することも可能となる。これにより、イベント映像の検索などが大幅に改善される。

【図面の簡単な説明】

【図1】 本発明自動インデックス発生装置の一実施形態をブロック図にて示している。

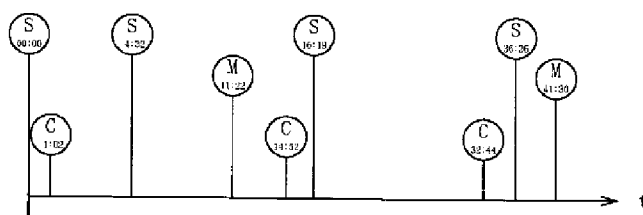
【図2】 図1中のインデックス発生・出力部から最終的に取り出されるインデックスの一例を示している。

【図3】 インデックス付与装置によりインデックスが付与された映像のみを取り出してサッカー番組のダイジェスト版を制作する場合を示している。

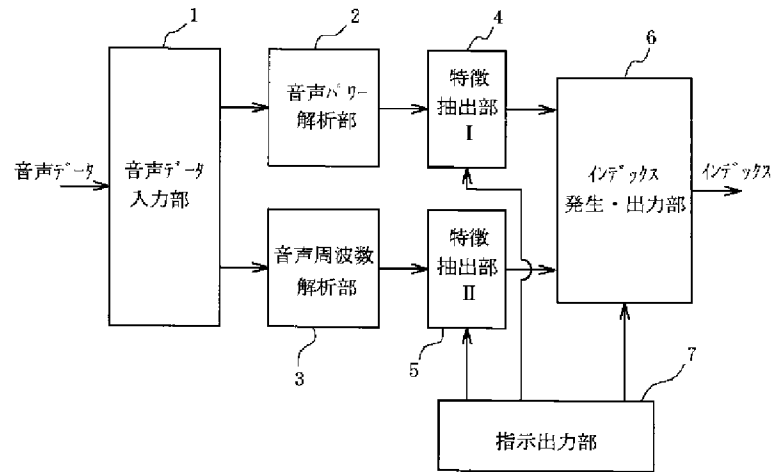
【符号の説明】

- 1 音声データ入力部
- 2 音声パワー解析部
- 3 音声周波数解析部
- 4 特徴抽出部I
- 5 特徴抽出部II
- 6 インデックス発生・出力部
- 7 指示出力部

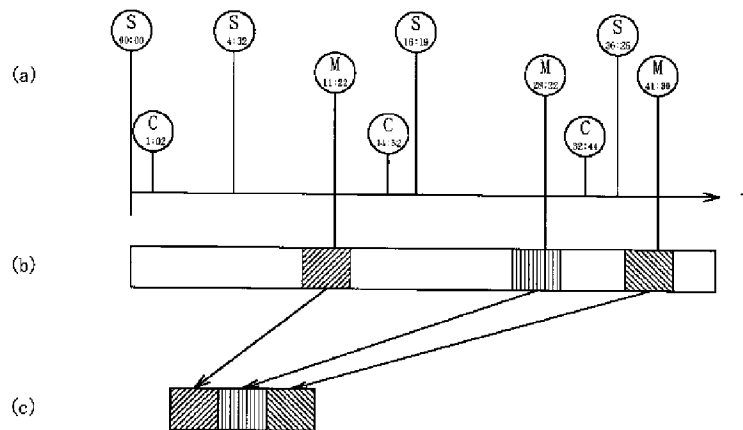
【図2】



【図1】



【図3】



フロントページの続き

(51) Int. Cl. ⁷	識別記号	F I	テーマコード [*] (参考)
G 1 0 L 15/00		G 1 0 L 3/00	5 3 1 N
G 1 1 B 27/36			5 5 1 G
H 0 4 N 5/91		H 0 4 N 5/91	C

(72) 発明者 住吉 英樹
東京都世田谷区砧 1 丁目 10 番 11 号 日本放送協会 放送技術研究所内

F ターム (参考) 5B075 ND14 NK39 PR01 UU35
5C053 FA14 GB06 GB11 HA29 JA03
JA12 KA14 LA06
5D015 AA06 CC01 CC14 HH03 HH23
KK00
5D077 AA01 CA02 CB07 DC16 DC21
DC40 DE01 DE11 DF01
5D087 AA02 AA22 AA24 CC02 CC27
DB02 DB03